Research Note: Adding Digit Vectors

Peter Hellekalek*

September 18, 2012

Abstract

In this paper, we study the different possibilities to add two vectors of digits of a given length m. Our results show that there are at least 2^{m-1} different additions of such vectors, while there exist only two *types* of addition that we may employ, addition with carry and addition without carry. The proofs of our results are elementary.

1 Introduction

In this research note we investigate the different possibilities to add two digit vectors of the same length.

Addition of digit vectors, in particular addition of binary vectors, is employed in many algorithms. Prominent examples in applied cryptography are the block ciphers IDEA [11] and AES [1] and several stream ciphers. We refer the reader to [12] and to the eSTREAM-project[1] for details on such ciphers.

The author's starting point into this question was the following observation. Any construction method for finite or infinite sequences of points is based on some arithmetical operations like addition or multiplication, on a suitable domain. It is most helpful if the algebraic structure underlying these operations is an abelian group. The choice of this group determines which function systems will be suitable for the analysis of a given sequence, because the construction method is intrinsically related to function systems, via the concept of the *dual group* (see Hewitt and Ross [10]). Different types of sequences require different types of function systems for their analysis. An example of such a suitable "match" between sequences and function systems in the theory of uniform distribution of sequences in the s-dimensional unit cube $[0,1)^s$ is given by Kronecker sequences or, in their discrete version, good lattice points, and the trigonometric functions. This construction method is based on *addition modulo one* (see Niederreiter [13, Ch. 5] and Sloan and Joe [16]). A second example are digital nets and sequences and, as appropriate function system, the Walsh functions. Here, *addition without carry* of digit vectors comes into play (see Niederreiter [13, Ch. 4] and Dick and Pillichshammer [2]).

*Research supported by the University of Salzburg, projects P1884/5-2009 and P1884/4-2010.
2010 *Mathematics Subject Classification*: Primary 11T71; Secondary 94A60.
Key words and phrases: XOR, addition, finite abelian groups, block ciphers, stream ciphers

[1] http://www.ecrypt.eu.org/stream/

One important type of a digital sequence, the Halton sequence, can also be generated by *addition with carry*, the underlying group being the compact abelian group of b-adic integers. From the search for the appropriate function system in this case, the notion of the b-adic function system originated. This concept was developed in series of papers (see [4, 5, 6, 8, 7], and for a background in ergodic theory [3]).

These investigations led to the question if there are any other types of addition of digit vectors, because if not, then the Walsh functions in base b and the b-adic function system and their combination in a *hybrid function system* (see [6] for this notion) cover all possible cases of function systems associated with additions of digit vectors.

Our results below show that this is indeed the case: there are *only two types* of addition of digit vectors: addition without carry, which is also called XOR-addition, and addition with carry, which is also known as integer addition.

We exhibit that, for a given length m of the digit vectors with digits in some given integer base $b \geq 2$, there are at least 2^{m-1} different additions for such vectors. This large number may be increased considerably if we employ also automorphisms of suitable groups of residues.

Our reasoning is elementary. It is based on a classical theorem on finite abelian groups and on the notion of compositions of positive integers.

The ideas presented below might have applications in cryptography, for example in stream or block cipher algorithms. If the information which digits are added in which way in the enciphering scheme is kept secret, then this will add not only to *confusion*, but, as already used in IDEA, changing the type of addition also adds to *diffusion* (for these two notions, see [15]). Hence, breaking the cipher would be more difficult.

2 Addition of digit vectors

Let $b \geq 2$ be a fixed integer and let $\mathcal{A}_b = \{0, 1, \ldots, b-1\}$ denote the set of b-ary digits. For $m \in \mathbb{N}$, let $\mathcal{A}_b^m$ stand for the m-fold cartesian product of the set $\mathcal{A}_b$ with itself.

We will study the following question, mostly in the case $b = p$ a prime: What are the binary operations "+" on the set $\mathcal{A}_b^m$ of digit vectors such that the pair $(\mathcal{A}_b^m, +)$ is an abelian group?

Remark 2.1. In this paper, when we speak of an "addition on $\mathcal{A}_b^m$", we mean a binary operation "+" on the set $\mathcal{A}_b^m$ of digit vectors in base b such that the pair $(\mathcal{A}_b^m, +)$ is an abelian group.

The reader should note that the term "binary" has two different meanings in this paper, which will become clear from the context. A binary operation on a set G is a map from the cartesian product $G \times G$ into G. Referring to the representation of real numbers in base $b = 2$, the elements of the set $\mathcal{A}_2^m$ are called binary vectors, and for $m = 1$ one speaks of binary digits.

Let us consider the case $b = 2$ first. There are two well known examples for addition of digit vectors. One is addition *without* carry and the other is addition *with* carry.

For $n \in \mathbb{N}$, $n \geq 2$, let $\mathbb{Z}/n\mathbb{Z}$ denote the additive group of residue classes modulo n. We identify this cyclic group with the set of integers $\{0, 1, \ldots, n-1\}$ equipped with addition modulo n.

Example 2.2 (Addition without carry)**.** We identify $\mathcal{A}_2$ with $\mathbb{Z}/2\mathbb{Z}$. For $\mathbf{x}, \mathbf{y} \in \mathcal{A}_2^m$, $\mathbf{x} = (x_0, \ldots, x_{m-1})$ and $\mathbf{y} = (y_0, \ldots, y_{m-1})$, we define

$$\mathbf{x} + \mathbf{y} = (x_0 \oplus y_0, \ldots, x_{m-1} \oplus y_{m-1}),$$

where '$\oplus$' denotes addition on $\mathbb{Z}/2\mathbb{Z}$, $0 \oplus 0 = 1 \oplus 1 = 0$, and $0 \oplus 1 = 1 \oplus 0 = 1$. The pair $(\mathcal{A}_2^m, +)$ is an abelian group. In fact, it is isomorphic to the product group $(\mathbb{Z}/2\mathbb{Z})^m$. We call this binary operation *addition without carry*, or XOR-*addition* of digit vectors.

Any nonnegative integer k, $0 \leq k < 2^m$, has a unique dyadic representation of the form $k = k_0 + k_1 2 + \cdots + k_{m-1} 2^{m-1}$ with digits $k_j \in \mathcal{A}_2$, $0 \leq j \leq m-1$.

Example 2.3 (Addition with carry)**.** We identify $\mathcal{A}_2^m$ with the group $\mathbb{Z}/2^m\mathbb{Z}$. For $\mathbf{x} \in \mathcal{A}_2^m$, $\mathbf{x} = (x_0, \ldots, x_{m-1})$, we define the map $\mathrm{int}_2 : \mathcal{A}_2^m \to \mathbb{Z}/2^m\mathbb{Z}$,

$$\mathrm{int}_2(\mathbf{x}) = x_0 + x_1 2 + \cdots + x_{m-1} 2^{m-1}.$$

Further, let $\mathrm{dig}_2 : \mathbb{Z}/2^m\mathbb{Z} \to \mathcal{A}_2^m$,

$$\mathrm{dig}_2(k) = (k_0, k_1, \ldots, k_{m-1}),$$

where $k = k_0 + k_1 2 + \cdots + k_{m-1} 2^{m-1}$ is the representation of k in base 2. Finally, for $\mathbf{x}, \mathbf{y} \in \mathcal{A}_2^m$, we define

$$\mathbf{x} + \mathbf{y} = \mathrm{dig}_2(\mathrm{int}_2(\mathbf{x}) + \mathrm{int}_2(\mathbf{y}) \pmod{2^m}).$$

With this binary operation the pair $(\mathcal{A}_2^m, +)$ is an abelian group. Clearly, it is isomorphic to the additive group $\mathbb{Z}/2^m\mathbb{Z}$. We call this type of binary operation *addition with carry* or *integer addition* of digit vectors.

For $m \geq 2$, our two examples are non-isomorphic groups, because one is cyclic and the other is not.

Apart from these two examples, are there any other possibilities to define addition on the set $\mathcal{A}_2^m$ of binary digit vectors?

From the Fundamental Theorem for Finite Abelian Groups (see [9, Sec. 10]) we obtain the following corollary. In this context, a *partition* of a positive integer m is a finite sequence $(t_i)_{i=1}^r$, $r \in \mathbb{N}$, of positive integers with the two properties (i) $t_1 \geq t_2 \geq \cdots \geq t_r$, and (ii) $t_1 + t_2 + \cdots + t_r = m$.

Corollary 2.4. *The non-isomorphic groups of order 2^m, $m \in \mathbb{N}$, are given by the product groups*

$$(\mathbb{Z}/2^{t_1}\mathbb{Z}) \times (\mathbb{Z}/2^{t_2}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/2^{t_r}\mathbb{Z}),$$

where $(t_i)_{i=1}^r$ is a partition of m.

Hence, in view of Corollary 2.4, an addition on the set $\mathcal{A}_2^m$ is defined if we put

$$(\mathcal{A}_2^m, +) \cong (\mathbb{Z}/2^{t_1}\mathbb{Z}) \times (\mathbb{Z}/2^{t_2}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/2^{t_r}\mathbb{Z}), \tag{1}$$

where $m = t_1 + t_2 + \cdots + t_r$ is a partition of m. Here, the symbol "$\cong$" denotes that the two groups are isomorphic.

As a consequence, there are at least as many possibilities to define addition on the set $\mathcal{A}_2^m$ of binary digit vectors of length m, as there are different partitions of the integer m.

From the structure of the factors in (1) we obtain the following information.

Corollary 2.5. *The only two types of binary operations on (sub)vectors of digits that may appear in the group law of the abelian group $(\mathcal{A}_2^m, +)$ are the following:*

- *addition given by finite product groups of the form $(\mathbb{Z}/2\mathbb{Z}) \times \cdots \times (\mathbb{Z}/2\mathbb{Z})$, which is what we have called* XOR*-addition, or*
- *addition in groups of residues of the form $\mathbb{Z}/2^t\mathbb{Z}$, $t \geq 2$, which we have called integer addition.*

Denote the number of different partitions of m by $P(m)$. We refer to the monograph [14, Ch. 2.5.1] for details on the partition function P like tables, or for results on its asymptotic behavior.

For example, if $m = 8$, then there are $P(8) = 22$ non-isomorphic groups of order 2^8, like the groups $\mathbb{Z}/2^8\mathbb{Z}$, $(\mathbb{Z}/2^7\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$, $(\mathbb{Z}/2^6\mathbb{Z}) \times (\mathbb{Z}/2^2\mathbb{Z})$, and so on. Among these 22 non-isomorphic groups of order 2^8, let us choose for illustration the group

$$(\mathbb{Z}/2^3\mathbb{Z}) \times (\mathbb{Z}/2^2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})^3.$$

What addition on $\mathcal{A}_2^8$ does this group induce? Addition of two bytes $\mathbf{x} = (x_0, \ldots, x_7)$ and $\mathbf{y} = (y_0, \ldots, y_7)$ is carried out as follows. The first three bits of $\mathbf{x}$ and $\mathbf{y}$ are interpreted as the three binary digits of two integers in the range $\{0, 1, \ldots, 7\}$. These two integers are added, the resulting integer is reduced modulo 2^3, which gives an integer in the range from 0 to 7, and the three binary digits of this integer give the first three digits of the sum $\mathbf{x} + \mathbf{y}$. In other words, for the first three bits, we carry out addition in the group $\mathbb{Z}/2^3\mathbb{Z}$ of residue classes modulo 2^3. The same procedure, which we have called integer addition, is applied to the next two bits. For the last three bits, the digit vectors (x_5, x_6, x_7) and (y_5, y_6, y_7) are XOR-ed, because we have to perform addition in the group $(\mathbb{Z}/2\mathbb{Z})^3$.

Observe that this is not the complete answer to our question. Our question concerned the different possibilities to add two binary vectors of length m. In the definition of such an addition, the position of each digit matters, whereas in Corollary 2.4 the order of the factors does not. The two groups

$$(\mathbb{Z}/2^3\mathbb{Z}) \times (\mathbb{Z}/2^2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})^3$$

and

$$(\mathbb{Z}/2\mathbb{Z})^3 \times (\mathbb{Z}/2^2\mathbb{Z}) \times (\mathbb{Z}/2^3\mathbb{Z})$$

are isomorphic, but they induce different additions on $\mathcal{A}_2^8$. In the first case, the first three digits of $\mathbf{x} + \mathbf{y}$ are computed via integer addition, in the second case these three digits are computed by the XOR-operation.

In terms of representing a number $m \in \mathbb{N}$, this means we consider the representation $8 = 3+2+1+1+1$ to be different from $8 = 1+1+1+2+3$, because they induce different additions on $\mathcal{A}_2^8$. This leads us to the following definition.

Definition 2.6. A *composition* of a positive integer m is a finite sequence of positive integers $(t_i)_{i=1}^r$, $r \in \mathbb{N}$, with the property $m = t_1 + t_2 + \cdots + t_r$.

Two such sequences which differ in the order of their summands are deemed to be different compositions, while they would be considered to be the same partition of m. The following result and its nice proof are well known.

Lemma 2.7. *Let $C(m)$ denote the number of different compositions of $m \in \mathbb{N}$. Then*

$$C(m) = 2^{m-1}.$$

Proof. The case $m = 1$ is trivial. Let $m \geq 2$. In the scheme

$$1\Box 1\Box \ldots \Box 1\Box 1,$$

of m 1's and $m-1$ boxes, we may replace every box either by a plus sign or by a comma. A different choice for each of the $m-1$ boxes leads to a different composition of m. □

We may summarize our findings as follows.

Theorem 2.8. *For the set $\mathcal{A}_2$ of binary digits and for $m \in \mathbb{N}$, the following holds for all binary operations "+" on the set $\mathcal{A}_2^m$ such that the pair $(\mathcal{A}_2^m, +)$ forms an abelian group:*

1. *There are only two types of addition of (sub)vectors of binary digits, addition without carry and addition with carry.*
2. *The number of different additions on $\mathcal{A}_2^m$ that arise from the compositions of m is equal to 2^{m-1}.*

The preceeding arguments may be generalized directly to the case of an arbitrary prime base p instead of base 2.

Corollary 2.9. *Let p be a prime. Then there exist only two types of addition for vectors of p-ary digits, addition without carry, which corresponds to addition on finite product groups of the form $(\mathbb{Z}/p\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p\mathbb{Z})$, and addition with carry, which corresponds to groups of the form $\mathbb{Z}/p^t\mathbb{Z}$, $t \geq 2$. Further, for every $m \in \mathbb{N}$, the number of additions on $\mathcal{A}_p^m$ that arise from the compositions of m is equal to 2^{m-1}.*

In case of a composite integer base b, $b \geq 2$, the situation is more complicated, because from the factorization of b into distinct prime powers many 'small' cyclic groups arise in the Fundamental Theorem for Finite Abelian Groups that cannot directly be related to operations on the b-adic digits, as we did above. One will have to use the Chinese Remainder Theorem to treat these cases. We omit these technical details because they do not contribute any new aspects to our investigation.

Clearly, every composition of the positive integer m defines an addition on $\mathcal{A}_b^m$, by simply following the recipes given above. Hence, the number of different binary operations "+" on $\mathcal{A}_b^m$ such that the pair $(\mathcal{A}_b^m, +)$ is a abelian group is at least 2^{m-1}.

In detail, the composition $m = t_1 + \cdots + t_r$ gives rise to the addition on $\mathcal{A}_b^m$ defined by the following product group:

$$(\mathcal{A}_b^m, +) \cong (\mathbb{Z}/b^{t_1}\mathbb{Z}) \times (\mathbb{Z}/b^{t_2}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/b^{t_r}\mathbb{Z}). \qquad (2)$$

3 Combination with automorphisms

In cryptographic applications of the ideas above, for example in stream ciphers, the information which of the m digits are added by XOR-addition and which by integer addition might become part of the key in the encryption scheme.

We will increase the key space considerably by the following idea. Suppose that we have chosen the composition $m = t_1 + \cdots + t_r$ of m. Hence, we obtain the group law on $\mathcal{A}_b^m$ from the product group given in (2). For a given factor $\mathbb{Z}/b^t\mathbb{Z}$ of this product we may combine integer addition with an arbitrary automorphism σ of the group $\mathbb{Z}/b^t\mathbb{Z}$ as follows. For $\mathbf{x}, \mathbf{y} \in \mathcal{A}_b^t$, $\mathbf{x} = (x_0, \ldots, x_{t-1})$ and $\mathbf{y} = (y_0, \ldots, y_{t-1})$, these t digits are added by the law

$$\mathbf{x} + \mathbf{y} = \mathrm{dig}_b\left(\sigma(\mathrm{int}_b(\mathbf{x})) + \sigma(\mathrm{int}_b(\mathbf{y})) \pmod{b^t}\right). \qquad (3)$$

Lemma 3.1. *There are $\varphi(b^t)$ different ways to define the addition in (3).*

Proof. The following reasoning is standard. Let σ be an homomorphism of the additive group $\mathbb{Z}/b^t\mathbb{Z}$ into itself. Then $\sigma(a) = a\sigma(1)$ for all $a \in \mathbb{Z}/b^t\mathbb{Z}$. Hence, σ is an automorphism if and only if $(a, b^t) = 1$. In other words, a belongs to the (multiplicative) group of prime residues $(\mathbb{Z}/b^t\mathbb{Z})^*$ modulo b^t, which has $\varphi(b^t)$ elements. □

Theorem 3.2. *Let p be a prime. Then, for any $m \in \mathbb{N}$, the compositions of m and the automorphisms of the associated groups of residues generate*

$$(p-1)(2p-1)^{m-1}$$

different additions on $\mathcal{A}_p^m$, i.e., binary operations "+" such that the pair $(\mathcal{A}_p^m, +)$ is an abelian group.

Proof. For a given composition $m = t_1 + \cdots + t_r$ of m into r components, $1 \leq r \leq m$, we obtain the group law from

$$(\mathcal{A}_p^m, +) \cong (\mathbb{Z}/p^{t_1}\mathbb{Z}) \times (\mathbb{Z}/p^{t_2}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p^{t_r}\mathbb{Z}). \qquad (4)$$

Due to Lemma 3.1, this group product allows

$$p^m \left(\frac{p-1}{p}\right)^r$$

automorphisms.

The number of compositions of m with r terms is equal to the number of possibilities to chose $r-1$ from the available $m-1$ places to put a comma in the proof of Lemma 2.7. For this reason, there exist

$$\binom{m-1}{r-1}$$

compositions of m into r terms.

Hence, the total number of different additions on $\mathcal{A}_p^m$ that stem from compositions of m and the associated automorphisms is given by

$$p^m \sum_{r=1}^{m} \binom{m-1}{r-1} \left(\frac{p-1}{p}\right)^r = (p-1)(2p-1)^{m-1}.$$

□

Example 3.3. Let $p=2$ and $m=8$. There are $2^7=128$ different additions on $\mathcal{A}_2^8$ arising from the 128 compositions of the number 8.

For the given composition $8 = 3+4+1$, there are $\varphi(2^3)=4$ different integer additions for the first three bits and $\varphi(2^4)=8$ for the next four bits, if we employ the combination of addition with automorphisms like in (3), and there is just one addition for the last bit. As a consequence, for this particular composition of $m=8$, there exists not only one addition of 8-bit dyadic vectors but there are 32 different additions available due to the combination with the 4 automorphisms of the factor $\mathbb{Z}/2^3\mathbb{Z}$ and the 8 automorphisms of $\mathbb{Z}/2^4\mathbb{Z}$.

Hence, if we allow automorphisms of the residue groups that appear as factors in the product group (4), then from Theorem 3.2 is follows that there are $3^7 = 2187$ different additions of 8-bit vectors available.

In the case of an arbitrary integer base b, the result is the following:

Theorem 3.4. *Let $b \geq 2$ be an integer. Then, for any $m \in \mathbb{N}$, the compositions of m and the automorphisms of the associated groups of residues generate*

$$b^m C_b (1+C_b)^{m-1}$$

different binary operations "+" on $\mathcal{A}_b^m$ such that the pair $(\mathcal{A}_b^m, +)$ is an abelian group. Here, the number C_b is defined as

$$C_b = \prod_{i=1}^{s} (1 - 1/p_i),$$

where $b = \prod_{i=1}^{s} p_i^{\alpha_i}$ is the factorization of b into distinct primes p_i, with $\alpha_i \in \mathbb{N}$, $1 \leq i \leq s$.

Proof. We translate the proof of Theorem 3.2 step by step from the case of a prime base p to the general base b. □

Acknowledgements

The author would like to thank Markus Neuhauser (RWTH Aachen) for a most helpful remark and the NUHAG research group at the University of Vienna for its hospitality during a stay at the Faculty of Mathematics that proved to be most fruitful.

References

[1] J. Daemen and V. Rijmen. *The Design of Rijndael.* Springer Verlag, New York, 2002.

[2] J. Dick and F. Pillichshammer. *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration.* Cambridge University Press, Cambridge, 2010.

[3] P. Grabner, P. Hellekalek, and P. Liardet. The dynamical point of view of low-discrepancy sequences. *Uniform Distribution Theory*, **7**:11–70, 2012.

[4] P. Hellekalek. A general discrepancy estimate based on p-adic arithmetics. *Acta Arith.*, **139**:117–129, 2009.

[5] P. Hellekalek. A notion of diaphony based on p-adic arithmetic. *Acta Arith.*, **145**:273–284, 2010.

[6] P. Hellekalek. Hybrid function systems in the theory of uniform distribution of sequences. In L. Plaskota and H. Wozniakowski, editors, *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (Warsaw, 2010)*, Lecture Notes in Statist., New York, 2012. Springer. To appear.

[7] P. Hellekalek and P. Kritzer. On the diaphony of some finite hybrid point sets. To appear in Acta Arith., 2012.

[8] P. Hellekalek and H. Niederreiter. Constructions of uniformly distributed sequences using the b-adic method. *Uniform Distribution Theory*, **6**:185–200, 2011.

[9] I.N. Herstein. *Abstract Algebra.* Wiley, New York, 3rd edition, 1999.

[10] Edwin Hewitt and Kenneth A. Ross. *Abstract Harmonic Analysis. Vol. I*, volume 115 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 1979.

[11] Xuejia Lai and James L. Massey. A proposal for a new block encryption standard. In *Advances in Cryptology—EUROCRYPT '90 (Aarhus, 1990)*, volume 473 of *Lecture Notes in Comput. Sci.*, pages 389–404. Springer, Berlin, 1991.

[12] A. J. Menezes, P. C. van Oorschot, and S.A. Vanstone. *Handbook of Applied Cryptography.* CRC Press, Boca Raton, 1997.

[13] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods.* SIAM, Philadelphia, 1992.

[14] Kenneth H. Rosen, John G. Michaels, Jonathan L. Gross, Jerrold W. Grossman, and Douglas R. Shier, editors. *Handbook of Discrete and Combinatorial Mathematics.* CRC Press, Boca Raton, FL, 2000.

[15] C. E. Shannon. Communication theory of secrecy systems. *Bell Sys. Tech. J.*, 28:656–715, 1949.

[16] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration.* Clarendon Press, Oxford, 1994.

Author's address:
Peter Hellekalek, Fachbereich Mathematik, Universität Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria
E-mail: peter.hellekalek@sbg.ac.at